

DOCUMENT RESUME

ED 145 691

FL 008 947

AUTHOR Dunckley, Candida J. Lutes; Radtke, Robert C.
TITLE Semantic Features, Perceptual Expectations, and Frequency as Factors in the Learning of Polar Spatial Adjective Concepts.
PUB DATE Mar 77
NOTE 42p.; Paper presented at the meeting of the Society for Research in Child Development (New Orleans, Louisiana, March, 1977)
EDRS PRICE MF-\$0.83 HC-\$2.06 Plus Postage.
DESCRIPTORS *Adjectives; Cognitive Processes; College Students; *Concept Formation; English; *Language Development; *Language Research; Learning Processes; *Learning Theories; Linguistic Theory; Preschool Children; Psycholinguistics; *Semantics; Visual Perception; Vocabulary Development; Word Frequency
IDENTIFIERS *Word Learning Theory

ABSTRACT

Two semantic theories of word learning, a perceptual complexity hypothesis (H. Clark, 1970) and a quantitative complexity hypothesis (E. Clark, 1972) were tested by teaching 24 preschoolers and 16 college students CVC labels for five polar spatial adjective concepts having single word representations in English, and for three having no direct translation into English. Neither theory was supported: polarity, verticality, and number of features were inconsistently related to speed of learning. The failure to replicate earlier studies is attributed to the fact that conceptual dimensions were unconfounded in the present experiment. At the end of the study, the preschoolers' understanding of the real English words was assessed in a multi-trial situation. The likelihood that any given concept would be responded to correctly was reasonably predicted by the frequency of occurrence of that word in the language environment. (Author)

* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *

ED145691

FL008947

Semantic Features, Perceptual Expectations, and Frequency
as Factors in the Learning of Polar Spatial Adjective Concepts

Candida J. Lutes Dunckley

Robert C. Radtke

Texas A&M University

Southern Illinois University, Carbondale

Paper presented at the 1977 meeting of the society for Research in
Child Development, New Orleans.

PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Candida J. L. Dunckley

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC) AND
THE ERIC SYSTEM CONTRACTORS.

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRO-
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN-
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRESENT
OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY.

Abstract

Two semantic theories of word learning, a perceptual complexity hypothesis (H. Clark, 1970) and a quantitative complexity hypothesis (E. Clark, 1972) were tested by teaching preschoolers and college students CVC labels for five polar spatial adjective concepts having single word representations in English, and for three having no direct translation into English. Neither theory was supported: polarity, verticality, and number of features were inconsistently related to speed of learning. The failure to replicate earlier studies is attributed to the fact that conceptual dimensions were unconfounded in the present experiment. At the end of the study, the preschoolers' understanding of the real English words was assessed in a multi-trial situation. The likelihood that any given concept would be responded to correctly was reasonably predicted by the frequency of occurrence of that word in the language environment.

Semantic Features, Perceptual Expectations, and Frequency
as Factors in the Learning of Polar Spatial Adjective Concepts

Candida J. Lutes Dunckley

Robert C. Radtke

Texas A&M University

Southern Illinois University, Carbondale

Two recent theories concerning how a word is learned stress the importance of the semantic features(i.e., elements of meaning) of which each word is composed. The first, a perceptual complexity theory (H. Clark, 1970, 1973), includes the suggestion that certain semantic features (e.g., (+Pol)) are more salient to young children than are other features (e.g., (-Pol)), thereby making words that are defined by the more salient features easier to learn. The second, a quantitative complexity theory (E. Clark, 1972, 1973a, 1973b, 1974), is based on the belief that the number of semantic features that are needed to adequately define a word will contribute to the ease or difficulty with which that word is learned. While a number of studies have been published apparently supporting each of these theories, none has adequately controlled for possible biasing effects of frequency of occurrence of a word in the language environment of a given learner, nor has any adequately tested for the learner's full understanding of a word before crediting him with "knowing" that word. The present study was an attempt to deal with these two criticisms.

E. Clark (1973a) was one of the first to present data relevant to H. Clark's perceptual complexity theory. Concerned with children's understanding of the locative prepositions in, on, and under, she showed that the visual stimulus configuration was more likely to determine a young child's response to a verbal request than was the particular preposition used, and second, that the preposition in, which can most often be correctly predicted by the environmental configuration, was linguistically understood at an earlier age than was either on or under. These data were used to show that perceptual expectancies do govern the order of word acquisition. It may be, however, that perceptual expectancies simply determine response only when a word is not understood, and that other factors (possibly frequency of occurrence) determine the sequence of learning. As

Huttenlocher (1974) points out, there may be an error in crediting the child with having learned in simply because he shows a high percentage of correct responses to requests using the word in--the real test of understanding would require him to differentiate in from other words.

Other studies apparently supporting H. Clark's perceptual complexity theory center on children's understanding of the polar adjectives more and less. Balfour (1968), Donaldson and Hales (1970), and Palermo (1973, 1974) have shown that children make more correct behavioral responses to questions utilizing the adjective more than they do to those using less. While this has been interpreted as showing that children understand more before less, in support of H. Clark's prediction that positive polar items are perceptually easier to deal with and thus learned earlier than negative polar items, Huttenlocher's (1974) criticism is applicable here as well. That is, although children make correct responses to requests about more at the same age that they are selecting stimuli representing more when asked about less, they may do this either because they do understand more and do not understand less, or else because they understand neither term and simply believe that larger items make better choices. There is no way to differentiate these explanations with the kind of data collected by the researchers cited above, although a more recent study (Kavanaugh, 1976) supports the idea that the choices available in the task rather than the child's semantic understanding of the words may be responsible for their apparent more-less confusion. A third explanation that needs to be considered as well is that even if children understand more at an earlier age than they comprehend less, it may not be due to the polarity, but may instead be due to frequency differences between more and less in the English Language.

One study (Klatzky, Clark, and Macken, 1973) did attempt to separate the effects of frequency from those of polarity by teaching children CVC labels for polar adjectives rather than assessing their knowledge of the actual English words. The researchers compared the number of trials to criterion for positive polar spatial adjectives (big, tall, high, and long) to the number of trials to criterion

for the negative concepts (small, short, low, and short) and found that positive CVCs were learned more quickly than negative CVCs. Although Klatzky *et. al.* claimed support for the perceptual complexity theory, their data are also subject to the criticism that children may have a bias to select larger things, with response bias rather than an understanding of the positive concepts thereby accounting for the higher percent of correct responses to the positive polar items. An inspection of their stimuli shows that all four stimulus sets could effectively be differentiated by applying the concept big-small, so that a single response bias of this nature could account for the outcome of all four data sets. Furthermore, their manipulation of substituting CVCs for English concepts did not eliminate the possibility that children were mediating from already familiar concepts or labels to the novel CVC names. To correct for this, some of the concepts tested in the present study have no translation into English.

Additional researchers (Brewer and Stone, 1975; Kuczaj and Maratsos, 1975) have been unable to find differences in the speed of acquisition of positive and negative polar labels. Because neither of the above studies was directed to discovering the possible influences of response bias or frequency, it is still not certain whether their failure to replicate the polarity effect should be attributed to characteristics peculiar to their stimuli and manner of questioning, or to the possibility that H. Clark's polarity prediction is in error.

Research directed toward E. Clark's quantitative complexity theory (E. Clark, 1972; Nelson and Benedict, 1974; Haviland and Clark, 1974; Maratsos, 1974; Brewer and Stone, 1975) is also subject to more than one interpretation. Of all of the studies listed above, only one (Maratsos, 1974) required the children to show a differentiated knowledge of specific lexical items, and none separated the potential effects of prior exposure to a word from number of semantic features. Furthermore, most of the studies have differed in their manner of determining whether or not a child knows the meaning of a word.

E. Clark (1972), relying partially on Bierwisch's (1967) analysis of the number of semantic features contained by each of several polar spatial adjective concepts,

found that children could provide the antonym for a word with fewer semantic features prior to the time that they knew antonyms for words with a greater number of features. The ability to provide an antonym was taken as evidence that the child "knew" that concept. However, children in her study were not required to show whether they knew when one of the adjectives but not the others would be used. Therefore, it is uncertain that these children should be credited with a "knowledge" that paralleled number of features. Perhaps more important, the number of features contained by the adjectives in her list was confounded perfectly with the frequency of occurrence of these words in American English (see frequency norms in Kučera and Francis, 1967; Howes, 1966). Therefore, when the acquisition order that she predicted was obtained, it might have been due to frequency; it might have been due to number of features; or it might have been due to the fact that she did not test for full acquisition of the concepts, which conceivably could have produced a different order of acquisition. Thus one aspect of the present study was to separate frequency of occurrence from number of features, and a second was to require the language-learner to differentiate each spatial dimension from every other before concluding that he had "acquired" that concept.

Nelson and Benedict (1974) and Haviland and Clark (1974) also presented word acquisition data that correlated with number of features. The former researchers used reaction time as a measure of cognitive complexity, and found that it correlated positively with number of semantic features. The latter required children to provide verbal definitions of kinship terms, and found that their ability to define a term also varied with the number of features in that term. It can be argued in each case, however, that first, the skills tapped probably go beyond a simple working knowledge of the words being tested, and that second, the number of features is again confounded with frequency in the language.

Both Haratsos (1974) and Brewer and Stone (1975) attempted to insure that children who were classed as understanding a concept such as tall-short could adequately differentiate that concept from some other, such as big-small. Brewer and Stone did not, however, test for full acquisition. On the other hand, by restricting himself to a small set of highly similar concepts (big-small,

tall-short, and high-low) Haratsos was able to completely cross the stimuli representing a given concept with the others. For example, by placing a toy turtle on a box beside a toy elephant, he could determine whether the child knew that big and high (or small and low) are not synonymous. That is, to respond correctly about high objects, the child could not consistently use a strategy of selecting big objects. In this respect Haratsos' study differed from all of the others that have been presented as tests of E. Clark's quantitative complexity theory. His technique of contrasts was used in the present study.

To remove the confounding of concepts with frequency in English, the present study included three concepts that have no direct translation into English. These concepts should not be considered in any way cognitively unnatural, however, since all represent spatial relationships that have direct translations in Japanese (K. E. Hochizuki, personal communication, University of California at Santa Barbara, July, 1975) and can be described as follows. The English depth concept, deep-shallow, can be divided into two more restrictive concepts, V-depth and H-depth, for which feature lists are provided in Table 1. V-depth refers to vertical indentations into horizontal surfaces, such as caves, bookshelves, picture frames with an inset surface for the picture, etc. The English concept big-small, which refers to any obvious size characteristic of an object, either two- or three-dimensional, can be made more restrictive as well. H-surface refers to the horizontal surface area of an object, and thus can be applied appropriately to table tops, floor space, ceiling area, counters, etc., but not to doors or walls (which are vertical). If the theoretical predictions concerning semantic features are correct, they should hold true for these "contrived" concepts as well as for those that do exist in simple form in English.

To determine whether or not a child understands a concept, he was required in the present study to show that he could use it correctly in all conditions--

that is, he was required to differentiate that concept from every other one tested. For example, he was required to show that he understood that big is not synonymous with tall, high, wide, deep, (+Pos) V-depth, (+Pos) H-depth, or (+Pos) H-surface. In an early part of this experiment the child was taught CVC labels for each of the eight concepts listed in Table 1, one concept at a time, using a discrimination task in which he had to choose one block from each pair in which the concept to be learned was contrasted with every other (eg., to select the big one on four successive trials when it was not simultaneously the tall, high, wide, or deep one respectively). At the end of the study, the child was asked to respond to questions about the actual English words that the CVC labels had earlier represented, again contrasting each concept with every other. The CVC discrimination task provided information on the difficulty of learning concepts that vary in number of semantic features, and that vary in type of feature (e.g., (+Pol) vs. (-Pol)); the final assessment of the child's understanding of the real English words permitted an estimate of the likelihood that the child could have mediated from the real English concept to the CVC label.

To determine whether response biases to select larger items might account for a differential ease of learning positive and negative polar CVC labels, children were asked to select a series of blocks (without feedback) from pairs prior to each discrimination task. Thus, if they typically were to select the positive item for a given concept when there was no externally imposed basis for choice, and if the trials to criterion for the positive pole of that concept were lower than for the negative pole, then it would remain possible that children believe that positive polar items make better choices, rather than that they actually understand the positive earlier than the negative pole.

Finally, to determine whether children who were taught CVC labels had actually learned concepts rather than simple automatic responses to stimuli, they

were asked to apply the labels to novel objects at the end of each discrimination training session. This was simply a check on the assumption that the experiment is, in fact, relevant to natural language knowledge.

In sum, the present study was designed to answer the following questions.

(1) H. Clark (1970, 1973) suggests that certain features of the environment are perceptually more salient to children than are other features, and that words representing the more salient features will be learned more easily. With frequency and response bias controlled for, are positive polar items in fact easier to learn than negative, and are concepts implying the vertical easier than those indicative of the horizontal? (2) E. Clark (1972, 1973a, 1973b, 1974) suggests that whenever one word has more features than a second, and that when every feature in the second word is also a feature of the first, then it should take longer to learn the first word. With frequency controlled for, and with full knowledge of each word ascertained, does number of features relate to difficulty of acquisition? (3) In that no study has systematically ascertained the likelihood that a child will correctly use any one of a set of actual English polar spatial adjectives, the present study assessed knowledge of real words and looked at the relationship between knowledge, number of semantic features, type of features, and frequency in the English language.

Method

Subjects

Twelve female and 12 male children attending a private preschool in Ventura, California who had been matched for age served as subjects. Their ages ranged from 30 to 65 months, with a mean of 53.7 months. The parents of 13 children had blue-collar professions, nine had white-collar professions, and data were unavailable on the remaining two families. All children who began the study completed the nine sessions, each taking roughly 15 to 25 minutes per

session. Testing took place in the preschool.

Eight female and eight male students from an introductory psychology course at Texas A & M University, ranging in age from 17 to 36 years, made up the adult sample. The mean age of the women was 20.8 years, and that of the men was 19.1. Each college student performed all tasks in a single session, which lasted about one hour and 15 minutes.

Materials

Wooden blocks, a variety of toys, and drawings of both the blocks and toys were used as testing materials in the present study. To loosely illustrate the nature of each of the eight polar spatial adjective concepts to the subjects, there were eight sets (five blocks each) of white wooden blocks that differed from each other along the dimension in question. That is, to illustrate the concept big-small there were five cubic blocks, ranging in equal increments from four to 14 cm. on a side. For tall-short there were five rectangular blocks, all with a base area of five x five cm., but ranging in equal height increments from four to nine cm. For wide-narrow the five blocks all had a depth of 12.5 cm. and a height of one cm., but ranged in width from two to eight cm. The blocks representing the concept high-low were each cubes four cm. on a side, and were placed on five shelves each eight cm. above the next. Blocks representing deep-shallow, V-depth, and H-depth were each cubes nine cm. on a side, with cylindrical holes drilled into them to a depth of from 0.3 cm. to 4.0 cm. For deep-shallow the location of the hole on the block was variable, while for V-depth each hole was in the top of the block, and for H-depth each hole went into the front surface. The blocks used to demonstrate the concept H-surface were each four cm. in height, with square top surfaces ranging from four x four to nine x nine cm.

In order to test for full acquisition of each of the eight concepts, an additional set of fifty-two pairs of white wooden blocks was constructed which

contrasted each conceptual dimension with every other (except that deep-shallow could not logically be contrasted with either H-depth or V-depth, reducing the number of potential pairs from 56 to 52). In half of the pairs, the block representing the positive pole of one dimension (e.g., big) represented the negative pole of the second (e.g., short), while in the remaining sets, like poles coincided. Therefore, reinforcement could be consistently associated with a single attribute only: stimulus dimensions were unconfounded.

The remaining stimulus materials were used to assess the responder's ability to apply the conceptual labels to real and pictured items. These included (1) drawings of the 52 block pairs described above--i.e., stimuli that differed in modality but not in form from the preceding set; (2) three-dimensional toy-sized objects that children are generally familiar with from their own daily experiences at home, preschool, or from children's books--i.e., stimuli that differed in form but not modality from the block pairs; and (3) drawings of the toy-sized objects--i.e., stimuli that differed both in form and in modality from the original set of blocks. Three representatives for each concept were randomly selected to test each subject. For big-small there were two rubber balls, two crayons, two file boxes, and two barrels from a stacking toy, with the two members of each pair differing in size. For tall-short, there were two horses, two barrels from a stacking toy, and two magnetic alphabet letters, this time with the two members of each pair differing in height. For deep-shallow there were two halves of a single barrel from the stacking toy, each half filled to a different level with clay, two bird feeders with clear plastic sides, each having a different interior depth, two dump trucks, and a sink and bathtub set. Wide-narrow was tested with the use of two dump trucks, two crayon boxes, and two books. High-low was tested with the use of books on two differently spaced shelves, two toy birds attached to a single bird feeder (one on the roof and the other on the bottom perch), and two alphabet letters placed at different

heights on a magnetic board. For H-surface there were two doll house tables, two books lying flat, two file boxes, and two barrels. For H-depth there were two toy refrigerators, two picture frames (each with the picture inset to a different depth), and two bird feeders (each with a different interior depth). Finally, for V-depth there were two dump trucks, a coffee pot and coffee cup, and a sink and bathtub set.

Procedure

Each child was tested individually, in a total of nine separate sessions that usually occurred on successive days, and in no case separated by more than two days. Each of the first eight sessions had, in order, (1) a preference test; (2) the demonstration of one concept using one of the eight sets of five blocks; (3) discrimination training in applying CVC labels to the two poles of that concept, using the subset of 14 from the 52 block pairs that was relevant to the concept in question; and (4) three tests for the ability to generalize the CVC labels to new stimuli (pictures of blocks, toys, and pictures of toys). In the ninth session each child was required to show the extent of his knowledge of the real English words representing all eight concepts.

Adults were also tested individually, but all portions of the task were conducted sequentially in a single session, and after the first two subjects, the ninth task was omitted for college students because it was tediously easy. No adults were tested for preference.

In all cases the subject and experimenter sat facing each other at a table. Stimulus materials not in use on any given trial were kept in an open suitcase and on shelves below the table, out of the subject's sight.

Only one of the eight dimensional concepts was used in each of the first eight sessions. The order in which the eight concepts became relevant varied for each subject. Sequences were random, with the restriction that each concept appear an

equal number of times in each of the eight positions (first, second, third, etc.) for each age and sex group.

In the third portion of each of the first eight sessions, children were taught CVC labels to represent the opposite poles of the concept in question. The eight pairs of CVC labels were selected from low association value lists (Hilgard, 1951), and were paired randomly with the eight conceptual dimensions with the following three restrictions. First, each pair of CVC items was associated with each concept within each age and sex group an equal number of times. Second, each CVC item within a pair was assigned to label the positive pole for half of the subjects and the negative for the remainder. And third, no subject was requested to learn the same CVC labels for more than one concept. This insured that any differences in the ease or difficulty with which different concepts or poles were learned could not be attributed to differences in CVC difficulty.

Preference Test. The experimenter placed the five-item demonstration set for a given concept on the table in front of the child, with the items arranged in order from one pole to the other. To determine whether the positive pole was the preferred response choice, each child was asked to point to the item that he liked best, and then to the item that he liked least (on half of the trials, the order of questioning was reversed). Pilot work suggested that younger children responded more readily to instructions phrased "Put your finger on the good (bad) one" and such wording was used whenever it seemed appropriate. Because three of the five response alternatives could not clearly be scored as polar, the two responses to the set of five Demonstration blocks were not scored; children responded to this set only to learn how to respond to the experimenter's request.

Following the presentation of the five-block Demonstration set for the first two preference trials, the 14 pairs (from the 52-pair Discrimination set) that

were relevant to the concept being tested were shown, one pair at a time. On half of the trials the child was requested to point to the item that he liked best, and on the remainder he was to indicate the member of the pair he liked the least. Thus, seven preferences and seven rejections were made prior to the onset of teaching the concept labels.

Concept Demonstration. Each child was told that he was going to play a game, the object of which was to learn some new words. The five Demonstration stimuli for the appropriate dimensional concept were placed on the table, arranged in ascending order from the most negative to the most positive example (with the negative to the left for half of the subjects and to the right for the remainder). The experimenter pointed to the negative and positive poles, one at a time, telling the child its CVC label. She then pointed to the central item in the display, remarking that it was in the middle, to impress the dimensional nature of the display on the child. The display was then rotated, keeping the five blocks in dimensional order, but reversing the poles from right to left. The polar labels were again provided by the experimenter. One at a time the experimenter pointed to the three non-extreme stimuli as a reference point, and she indicated that all blocks to one side of the reference point would receive one CVC label relative to that block, and that all blocks to the other side would receive the other CVC label. Finally, the display was rotated again, and each block was again labeled by the experimenter relative to one of the non-extreme reference points. The purpose of this demonstration was to quickly familiarize the child with the CVC labels and with gross aspects of the concept before teaching a more refined understanding of that concept. Following this demonstration the child was asked to point to instances of each CVC label relative to some reference point indicated by the experimenter. If the child was in error on any trial, he was corrected. This procedure was continued until the child could select appropriate items for both the positive and negative poles twice in succession on the Demonstration display.

It should be noted that this portion of the study does not constitute adequate training of a concept because each concept is at this point confounded to a large extent with defining features of the other concepts. In addition, the criterion of success is extremely weak.

Discrimination Training. Following the initial demonstration, the experimenter removed the Demonstration items and one at a time introduced the 14 pairs (from the larger set of 52) that were relevant to the concept being tested. As she introduced the first pair she stated, "I have more blocks here and want to know if you can tell me the words for them. Here are two blocks. Show me the one that is ____" (beginning with the positive item for half of the subjects and the negative item for the remainder). Only one label was requested on each trial, and the actual sequence of positive- and negative-label requests, as well as the specific stimulus pair presented, was determined for each concept for each child from a table of random numbers, with two restrictions. First, the same pole was not requested more than three times in succession; and second, none of the fourteen stimulus pairs relevant to a given concept were presented a second time until all pairs had been presented once. Verbal feedback was given on each trial. Training for each subject was continued until a criterion of fourteen consecutive correct responses per pole was attained.

Generalization Tests. The three sets of Generalization stimuli (toys, block pictures, toy pictures) were presented, one at a time, and in a different order for each of the concepts the subject had to learn. Across subjects in an age group, the three Generalization tests appeared in each possible order an equal number of times for each concept. For pictorial stimuli, the experimenter held up an illustration of a single pair of items and said "Point to the one that is ____" (with positive and negative CVC labels requested in random order). Each pair was shown twice, requesting a different pole each time, for a total of six generalization trials to that set of stimuli, and with three requests for each

CVC pole. For the actual toys, two items were set out at a time, with the same requests made as for pictorial stimuli. No feedback or corrections were given on any of the generalization trials.

English Equivalents. In the ninth session each child was shown, one pair at a time, the 26 pairs of Discrimination blocks in which one member represented the positive pole of the first concept but the negative pole of the second. The 26 pairs were presented in random order, four times through the set, resulting in 104 trials. On each trial only one label was requested; to choose correctly the child would have to differentiate the two concepts contrasted in that block pair. The labels requested were big, small, tall, short, high, low, wide, narrow, deep, shallow, biggest top, smallest top, deepest hole in top, shallowest hole in top, deepest hole in front, and shallowest hole in front. The percent of correct responses to each label was computed. Children were given no feedback in this portion of the test.

Data Analyses. In each type of data, H. Clark's perceptual complexity theory was tested by observing the main effect of, and interactions involving, the polarity feature as well as effects due to the presence or absence of the feature (+Vert). E. Clark's quantitative complexity theory was investigated in a broad way by looking for main effects and interactions involving the difficulty of the three- vs. four-feature concepts. More narrowly, it was investigated by comparing tall-short with V-depth; deep-shallow with both H-depth and V-depth; and tall-short with high-low. In each comparison the first member was expected to be easier than the second because the second was defined by adding one feature to the list of the first, with all other features being identical. In those contrasts involving H-depth and V-depth we would also expect a difference in difficulty simply on the basis of the fact that both are contrived concepts (i.e., they have no direct translation into English), and this test is therefore biased slightly in favor of E. Clark's hypothesis. Finally, the frequency theory was tested by contrasting the natural

with the contrived concepts, and by contrasting the two most frequent natural concepts (big-small and tall-short) with three least frequent (high-low, wide-narrow, and deep-shallow).

Results

Preference Test

In general, children's responses to the preference test indicated that for about half of the children on each concept there was no polar preference. Those children expressing a preference typically selected the positive item, giving some support to the notion that "bigger" things are viewed as better choices. However, because neither the number of positive nor of negative polar preferences for a single concept was predictive of any other dependent measure for that concept, the suspicion that apparent differences in concept-learning scores might reflect a polar response bias rather than superior learning of a single pole was greatly reduced.

Specifically, the number of times that a child selected the positive exemplar as most preferred, and the number of times that the negative was most preferred, were each subjected to a 2(age: above vs. below the median) x 2(sex) x 8(concept) mixed analysis of variance. No main effects or interactions were significant at the 0.05 level. The data were then re-analyzed by categorizing children in one of three ways: as preferring the positive pole (at least 11 consistent selections out of 14), as preferring the negative pole (at least 11 consistent selections out of 14), or as having no polar preference (between four and 10 selections of a single pole out of 14). The probability of making 11 or more consistent selections out of 14 when in fact choice is random is .057 (Heintraub, 1963). Roughly half of the children expressed some polar preference when tested for each concept, the greatest proportion selecting positive items rather than negative, $z = 6.307$, $p < .0001$. The proportion of children preferring one pole over the other was not significant for any single concept, however. Finally, preference scores were correlated (Pearson r) with all other dependent

measures of the study. Except that the number of rejections of the positive item is clearly predicted by the number of preferences, $r = -.787$, $z = -10.876$, $p < .0001$, none of these correlations reached significance at the .05 level.

Concept Demonstration

Data collected from this test were somewhat incidental to the focus of the study, in that the purpose of the Concept Demonstration was to speed the learning process rather than to assess the level of understanding of the subjects. However, because this portion of the experiment used a methodology similar to that of Klatzky, Clark, and Macken (1973), it was analyzed for polarity effects to determine whether the positive pole superiority was replicated. It was not; on those concepts in which a difference existed, the negative pole appeared to be easier. Concept differences were also investigated, but the outcome of this analysis should not be weighed too heavily because the response criterion was extremely weak. None-the-less, consistent with other dependent measures in the study, number of features was not predictive of scores, either in the gross sense of three vs. four features or in the narrow sense of contrasting pairs of concepts that differed with respect to only one feature.

With a criterion of only two consecutive correct responses for each polar label, very few errors were made by adults in this phase of the task, and with a mean of 0 in most cells, statistical analyses were inappropriate. For children, a $2(\text{age}) \times 2(\text{sex}) \times 2(\text{polarity}) \times 8(\text{concept})$ mixed analysis of variance of the number of errors made showed that children above the median age ($\bar{X} = .29$) did, as expected, make fewer errors than children below the median age ($\bar{X} = 0.60$), $F(1,20) = 6.3603$, $p < .025$. Furthermore, concept differences, $F(7,140) = 3.2042$, $p < .005$, and polarity effects, $F(1,20) = 8.0477$, $p < .025$, were significant but are qualified by a significant concept $\times$ polarity interaction, $F(7,140) = 3.0185$, $p < .01$. Collapsing over concepts, negative items ($\bar{X} = .356$) were easier than the positive ($\bar{X} = .614$). However, these polar effects are significant at the .05

level using orthogonal contrasts for only four of the eight concepts (deep-shallow, $F(1,184) = 4.51$, $p < .05$ wide-narrow, $F(1,184) = 7.80$, $p < .01$, high-low, $F(1,184) = 5.25$, $p < .05$, and H-surface, $F(1,184) = 5.51$, $p < .05$). Collapsing over polarity, a comparison of each concept with every other using the Tukey W.S.D. procedure (Miller, 1966, p. 92) showed that big-small was easier than wide-narrow, $q(140,8) = 4.62$, $p < .05$, high-low, $q(140,8) = 4.62$, $p < .05$, and H-surface, $q(140,8) = 4.99$, $p < .05$. No other differences between pairs were significant. Furthermore, planned orthogonal comparisons showed no significant differences for preschoolers due to the natural vs. contrived distinction (i.e., concepts with vs. without direct translation into English), $F(1,120) = .07$, due to the presence of three vs. four features, $F(1,120) = .79$, due to the feature (+Vert) vs. (-Vert), $F(1, 120) = .01$, nor to high- vs. low-frequency concepts, $F(1,120) = 1.14$.

Discrimination Task

The Discrimination Task differed from the Demonstration Task in that its stimulus dimensions were unconfounded (thus requiring the subject to attend to a specific concept in order to reach criterion), and in that the criterion was much more stringent (14 consecutive correct responses to each pole). The analysis of this data is therefore more relevant to each test of theory concerning word-learning. For adults, the only significant outcomes were that concepts utilizing the feature (+Vert) were learned more easily than were the others, and that the contrived concept H-surface was more difficult to learn than the others combined. For preschoolers, the only significant outcome was the relative greater ease of learning the positive than negative pole of the concepts high-low and V-depth.

Specifically, for adults a $2(\text{sex}) \times 8(\text{concept}) \times 2(\text{polarity})$ mixed analysis of variance revealed only the concept factor to be significant, $F(7,98) = 2.2227$, $p < .05$. Tukey W.S.D. testing, which is fairly conservative, revealed no significant differences between pairs of concept means. However,

planned orthogonal contrasts showed that the concepts implying verticality had the lowest error rates, $F(1,120) = 4.6284$, $p < .05$. Three-feature concepts did not differ from four-feature concepts, $F(1, 120) = .8328$, nor did natural differ from contrived, $F(1,120) = 1.975$, or high- from low-frequency concepts, $F(1,120) = .0723$. The most difficult concept, H-surface, was found to be more difficult than all others combined, $F(1,120) = 3.9330$, $p < .05$.

For preschoolers, a $2(\text{age}) \times 2(\text{sex}) \times 8(\text{concept}) \times 2(\text{polarity})$ mixed analysis of variance showed no significant effects except for the concept $\times$ polarity interaction, $F(7,140) = 5.2979$, $p < .001$. Orthogonal contrasts revealed that this is attributable to the smaller number of errors made to the positive than to the negative pole of high-low, $F(1,368) = 41.527$, $p < .01$, and of V-depth, $F(1,368) = 6.939$, $p < .01$. None of the planned orthogonal contrasts (natural vs. contrived, high- vs. low-frequency, and three- vs. four-feature concepts, all F values less than 1.0) reached significance at the .05 level.

Generalization Data

For adults, the incidence of errors was extremely low, and statistical analyses are therefore inappropriate. Preschoolers made more Generalization Task errors than did adults, although in an absolute sense they made few errors. There was a range from no errors (by children of any age) when asked to select real toys representing the concept H-surface, to a maximum of 1.67 errors (out of six trials) for children below the median age when asked to select drawings of blocks representing the concept V-depth. Across concepts, the mode was 0 errors out of six selections. Because children were able to apply the CVC labels appropriately to new stimuli, it is apparent that they were, in fact, learning linguistic concepts rather than specific stimulus-response connections.

A $2(\text{age}) \times 2(\text{sex}) \times 8(\text{concept}) \times 2(\text{polarity}) \times 3(\text{type of generalization items})$ mixed analysis of variance of preschoolers' error scores revealed a significant age effect, $F(1,20) = 7.0490$, $p < .025$, with children above the median age making fewer errors overall ($\bar{X} = .40$) than those below the median age ($\bar{X} = .71$). In addition, positive polar items ($\bar{X} = .58$) produced slightly more errors than

the negative ($\bar{X} = .53$), $F(1,20) = 6.008$, $p < .025$. The concept factor was also significant, $F(7, 140) = 2.2257$, $p < .05$, and Tukey H.S.D. contrasts between all pairs of means revealed that big-small, with the highest error rate, differs from the two concepts with the fewest errors, H-surface, $q(140,8) = 5.39$, and tall-short, $q(140,8) = 5.13$, at the .01 level; no other differences were significant. One other main effect, the type of generalization task, was also significant, $F(2,40) = 34.5667$, $p < .001$. The Tukey H.S.D. procedure showed that real toys ($\bar{X} = .333$) were easier to respond to correctly than were drawings of toys ($\bar{X} = .497$), $q(40,3) = 3.85$, $p < .05$; that drawings of toys were easier to respond to than were drawings of blocks ($\bar{X} = .826$), $q(40,2) = 7.70$, $p < .01$; and that real toys were easier to use than were drawings of blocks, $q(40,3) = 11.55$, $p < .01$.

Four interactions were significant. The sex x concept interaction, $F(7,140) = 2.6701$, $p < .025$, was broken down using orthogonal contrasts to show that females made more errors than did males (.05 level) on big-small, $F(1,1136) = 4.41$, and on V-depth, ($F(1,1136) = 4.41$, while males made more errors (.01 level) on H-depth, $F(1,1136) = 16.13$. The age x sex x concept interaction, $F(7,140) = 2.3262$, $p < .05$, indicates that the male superiority on the concept big-small occurs only in the older group, $F(1,1120) = 18.15$ $p < .01$, and that a female superiority on V-depth exists in the younger group, $F(1,1120) = 4.27$, $p < .05$. The age x generalization task interaction, $F(2,40) = 7.4016$, $p < .005$, is attributed to the finding that older children were superior to the younger (.01 level) when either the block drawings, $F(1,1146) = 44.346$, or the toy drawings, $F(1,1146) = 8.902$, were used, but not when real toys were presented, $F(1,1146) = 1.807$, n.s. Finally, the polarity x generalization task interaction, $F(2,280) = 5.1087$, $p < .025$, appeared to be due to the fact that positive items were more difficult when the task stimuli were drawings (either of blocks, $F(1,1146) = 12.1404$, $p < .01$, or of toys, $F(1,1146) = 5.6037$, $p < .05$) but not when real toys were presented,

$F(1,1146) = 2.8078$. It is interesting to note throughout this section of data that although in all three types of generalization tests the same conceptual knowledge was supposedly being tapped, scores differed from test to test. This points out the need to use more than one measure of concept understanding.

English Equivalents

The most important finding in this set of data is that although neither number of features nor polarity was predictive of a child's word knowledge when stimulus dimensions were unconfounded, frequency of occurrence in English was.

Each child responded to seven questions per pole per concept, except for the concepts deep-shallow, H-depth, and V-depth, because the latter three could not logically be contrasted. There were five questions concerning each pole of deep-shallow and six questions about each pole of the H-depth and V-depth concepts. Because there were unequal numbers of requests for each concept, the number of correct responses for each English label was converted to a percentage; these data are presented in Table 3.

A $2(\text{age}) \times 2(\text{sex}) \times 8(\text{concept}) \times 2(\text{polarity})$ mixed analysis of variance showed that older children ($\bar{X} = 79.25\%$) knew more words than the younger ($\bar{X} = 54.76\%$), $F(1,20) = 22.5983$, $p < .001$. In addition, the concept main effect was significant, $F(7,140) = 7.493$, $p < .001$, in that the least known concept, H-depth, had significantly more errors (.01 level), according to the Tukey W.S.D. test, than the three best known concepts, big-small, $q(140,8) = 7.61$, high-low, $q(140,8) = 6.24$, and deep-shallow, $q(140,8) = 6.07$. The concept with the second most errors, wide-narrow, was also significantly different (.01 level) from the three best known concepts, with $q(140,8) = 7.09$, 5.72 , and 5.56 respectively. The concept with the third highest error rate, H-surface, had more errors than big-small, $q(140,8) = 4.73$, $p < .05$; and tall-short, with the fourth highest error rate, also produced more errors than big-small, $q(140,8) = 4.67$, $p < .05$. Except for the concept deep-shallow, all of the concept means for

the five natural concepts were ranked in the order predicted by frequency. The scores of the contrived concepts are also ranked according to the frequency of occurrence in English of the words used to represent them in the Equivalents test.

Natural concepts ($\bar{X} = 70.16\%$) were better known than the contrived ($\bar{X} = 61.73\%$), $F(1,376) = 10.9596$, $p < .001$, and of the natural concepts, those of high frequency ($\bar{X} = 77.97\%$) were better known than the low frequency ($\bar{X} = 64.07\%$), $F(1,152) = 6.9742$, $p < .01$. The presence of the feature (+Vert) was not predictive, $F(1,152) = 1.8063$, nor was the three- vs. four-feature distinction, $F(1,152) = .0599$. In terms of concepts sharing a semantic field, i.e., the more narrow feature theory prediction, tall-short should have been easier than high-low and V-depth but was not; deep-shallow should have been easier than both H-depth and V-depth but was significantly easier than H-depth only.

Overall, positive labels ($\bar{X} = 70.89$) were better known than the negatives ($\bar{X} = 63.11$), $F(1,20) = 12.2744$, $p < .001$, but this is qualified by two significant interactions. First, the concept x polarity interaction, $F(7,140) = 4.0047$, $p < .001$, arose because the positive items were superior only for the concepts deep-shallow, $F(1,368) = 8.4102$, $p < .01$, high-low, $F(1,368) = 17.1743$, $p < .01$, H-depth, $F(1,368) = 4.0493$, $p < .05$, and V-depth, $F(1,368) = 9.5291$, $p < .01$. It should be noted that the direction of difference between poles of all concepts corresponds perfectly, except for tall-short, to frequency of use in the English language. The age x concept x polarity interaction was also significant, $F(7,140) = 2.2255$, $p < .05$. Subsequent orthogonal contrasts showed that while younger children in this study knew small (the more frequent member of the pair) better than big, $F(1,352) = 6.79$, $p < .05$, there was no polarity difference in the older group. Similarly, the positive (and also more frequently occurring) member of the pairs high-low, H-depth, and V-depth received a higher number of correct responses than did the negative member among the younger, but not older, children, $F(1,352) = 22.27$, $p < .01$; 5.18 , $p < .05$; 8.11 , $p < .01$, respectively. Finally, in the older, but not the younger group, deep was better known than

shallow, $F(1,352) = 5.28$, $p < .01$, and wide better known than narrow, $F(1,352) = 5.20$, $p < .05$; in each case the positive-pole superiority is predicted by frequency in the English language. Wide-narrow and deep-shallow are, overall, the least frequent of the natural concepts; therefore, polarity effects as a function of frequency might be expected to appear later than similar polarity effects in the more common concepts.

To determine whether children might have been mediating from familiar concepts to the nonsense syllables in the Demonstration and Discrimination portions of the task, their English Equivalents scores for each concept were correlated with the number of errors made in learning each concept during each of the earlier two tasks. None of these correlations (out of 32) were significant (.05 level) except for the correlation involving the positive pole of V-depth in the Demonstration task ($r = .49$, $p < .05$). This reduces the likelihood that their actual word knowledge had influenced their learning scores.

Discussion

Virtually no support was generated for H. Clark's (1970,1973) perceptual complexity hypothesis, i.e., that positive polar labels are easier to learn than the negative, and that spatial adjective concepts that refer to the vertical orientation are easier to learn than the non-vertical concepts. Specifically, in the Concept Demonstration, preschoolers tended to learn the negative pole more quickly in those instances where one pole was significantly easier than the second. In the Discrimination Task, the positive pole was superior for preschoolers for only two of the eight concepts tested, suggesting that the effect is not overly pervasive. For adults, concepts utilizing the feature (+Vert) were learned more easily than the others, but because this effect did not occur in the preschool sample, it is more likely that it arises somewhere during the course of development, rather than arises early to shape linguistic development as H. Clark suggests. On the Generalization Tasks, preschoolers made slightly

more errors to positive items than to the negative, contrary to Clark's position. In the English Equivalents Test, which assessed knowledge of real English words, neither polarity nor verticality was predictive of the likelihood that the child could use any given word correctly in a variety of situations. In that a child's response biases prior to Discrimination testing were not predictive of learning scores, the likelihood that response bias influenced learning score is greatly reduced.

The quantitative complexity theory of E. Clark (1972, 1973a, 1973b, 1974) fared no better than the perceptual complexity theory. Although in each portion of the study certain concepts were learned more efficiently than others, ease of learning did not relate to the number of features utilized by each concept. Neither the three- vs. four-feature distinction across concepts was predictive of learning score in any of the tests, nor did comparisons prove fruitful between the concepts tall-short and V-depth, between deep-shallow and both H-depth and V-depth, or between tall-short and high-low.

Although frequency of use in the English language was not predictive of concept learning score when CVC labels were used, it was reasonably predictive of knowledge of actual English words. The two exceptions to this were that deep-shallow was better known than predicted, and that the word tall received a higher percent of correct responses than short. It is not possible at this time to be certain whether these exceptions actually fail to match frequency, or whether the frequency predictions for children are inaccurate. Existing frequency norms typically reflect written language, which is particularly inappropriate for preschoolers. Although one of the frequency sources used in the present study (Howes, 1966) was based on actual speech, this "speech" was actually the uninterrupted monologues of a number of college students rather than dialogues involving preschoolers. Anecdotal evidence from the parents of children in the present study suggests that at least in the case of deep-shallow, frequency norms are inaccurate. The children tested live in a coastal area of Southern California and were tested in the summer; those parents who were asked verified that they had

indeed stressed the meaning of the words deep and shallow to their preschoolers. The current assumption made admittedly for practical purposes, that frequency norms are constant across age groups and geographic regions, should receive serious question and be put to empirical test.

Although the present study tends to refute both of the semantic theories tested, it suffers from the problem that the introduction of the Concept Demonstration prior to the totally trial-and-error Discrimination Task might have differentially affected Discrimination learning times for the different concepts. Therefore it may be important to repeat the experiment with the Demonstration omitted. However, because no semantic feature effect was evident even in the English Equivalents Task (which was not apparently influenced to any large extent by the preceding tasks), any feature effect that exists seems to be overridden by frequency of exposure, diminishing that factor's importance in understanding word-learning.

A logical problem with E. Clark's quantitative complexity theory is that it makes the a priori assumption that children's linguistic errors will always be errors of over-extension which must then gradually be narrowed as the child becomes cognizant of additional semantic features. There are, in fact, examples of errors of under-extension in the research literature (e.g., Reich, 1976) suggesting that a child may sometimes attend to numerous features and gradually eliminate attention to a part of the set.

Additionally, the process of defining different adjectives by constructing feature lists is itself somewhat suspect, in that different researchers do not agree as to what is the correct feature list for a given adjective. For example, although E. Clark (1972) stated that her feature lists were obtained from Bierwisch's (1967) analysis, inspection shows that their lists differed for big-small, long-short, thick-thin, and deep-shallow. Bierwisch considers big-small to be a three-feature item; Clark reduced it to two. Many of the feature lists provided by these authors can be argued against on the grounds that they do not

accurately provide the meanings of the words that they supposedly define. For example, Clark's definition of long-short includes the feature (-Vert), implying that these adjectives can be applied only to horizontal objects. In that we typically refer to dresses (vertical orientation) as long or short, the feature list must be inaccurate. Furthermore, Clark considers both tall-short and high-low to be three-feature concepts, and yet her feature lists show the same three features for both concepts, implying that tall-short and high-low are synonymous. Standard word usage refutes this; high-low refers to position in space, while tall-short refers to the extent of the object itself without reference to position in space, and so her feature list must be amended in that instance as well. For big-small, for which Clark has provided a list of two features, she gives no basis for deciding which spatial dimensions of an object to attend to when deciding whether that object is big or small. In reality, we do not randomly select dimensions in making this judgment--we select the functionally most important. For example, although big-small most often describes the overall volume (three dimensions) of an object, the "bigness" of items such as tables is usually assessed independent of height of the table or even of thickness of the table top. The height is much less important to the table's function than is its top surface area, and thus height may not be included in the size judgment. A Japanese dining table elevated only a foot off the floor is likely to be considered bigger than a European or American table with a smaller surface that is three to four feet off the floor. A butcher block is likely to be thought of as smaller than a standard dining table, when assessing it as a table rather than as a butcher block. Despite the foregoing arguments, however, it remains possible that semantic features are relevant to a child's learning of words, but that these features must be determined empirically rather than philosophically.

In sum, the present study, which carefully unconfounded the conceptual dimensions being investigated, provides no support for either H. Clark's

perceptual complexity hypothesis, nor for E. Clark's quantitative complexity hypothesis. It does suggest that a good predictor of whether or not a preschooler will understand the meaning of several polar spatial adjectives is simply the frequency with which each adjective occurs in his language environment.

References

- Bierwisch, H. Some semantic universals of German adjectivals. Foundations of Language, 1967, 3, 1-36.
- Brewer, H. F., & Stone, J. B. Acquisition of spatial antonym pairs. Journal of Experimental Child Psychology, 1975, 19, 299-307.
- Clark, E. V. On the child's acquisition of antonyms in two semantic fields. Journal of Verbal Learning and Verbal Behavior, 1972, 11, 750-753.
- Clark, E. V. How children describe time and order. In C.A. Ferguson and D.I. Slobin (Eds.), Studies of Child Language Development. New York: Holt, Rinehart, and Winston, 1973a. Pp. 585 - 606.
- Clark, E. V. Nonlinguistic strategies and the acquisition of word meanings. Cognition, 1973b, 2(2), 161-182.
- Clark, E. V. Some aspects of the conceptual basis for first language acquisition. In R. L. Schiefelbusch and L.L. Lloyd (Eds.), Language Perspectives--Acquisition, Retardation, and Intervention Baltimore, Md.: U. Pk. Press, 1974.
- Clark, H. H. The primitive nature of children's relational concepts. In J.R. Hayes (Ed.), Cognition and the Development of Language, New York: Wiley, 1970. Pp. 269-278.
- Clark, H. H. Space, time, semantics, and the child. In T. Moore (Ed.), Cognitive Development and the Acquisition of Language. New York: Academic Press, 1973. Pp. 27-63.
- Donaldson, H., and Balfour, G. Less is more: A study of language comprehension in children. British Journal of Psychology, 1968, 59(4)- 461-471.
- Donaldson, H., and Wales, R. J. On the acquisition of some relational terms. In J.R. Hayes (Ed.), Cognition and the Development of Language. New York: Wiley, 1970.

- Haviland, S. E., and Clark, E. V. This man's father is my father's son': A study of the acquisition of English kin terms. Journal of Child Language, 1974, 1, 23-47.
- Hilgard, E. R. Methods and procedures in the study of learning. In S. S. Stevens (Ed.) Handbook of Experimental Psychology. New York: Wiley, 1951.
- Hoves, D. A word count of spoken English. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 572-604.
- Huttenlocher, J. The origins of language comprehension. In R. L. Solso (Ed.), Theories in Cognitive Psychology: The Loyola Symposium. New York: Wiley, 1974. Pp. 331-368.
- Kavanaugh, R. D. On the synonymy of more and less: Comments on a methodology. Child Development, 1976, 47, 885-887.
- Klatzky, R. L., Clark, E. V., and Macken, M. Asymmetries in the acquisition of polar adjectives: Linguistic or conceptual? Journal of Experimental Child Psychology, 1973, 16, 32-46.
- Kucera, H., and Francis, W. H. Computational Analysis of Present Day English. Providence, Rhode Island: Brown University Press, 1967.
- Kuczaj, S. A., II, and Maratsos, M. P. On the acquisition of front, back, and side. Child Development, 1974, 45, 202-210.
- Maratsos, M. P. When is a high thing the big one? Developmental Psychology, 1974, 10, 367-375.
- Miller, R. G., Jr. Simultaneous Statistical Inference. New York: McGraw-Hill, 1966.
- Nelson, K., and Benedict, H. The comprehension of relative, absolute, and contrastive adjectives by young children. Journal of Psycholinguistic Research, 1974, 3(4), 33-341.
- Palermo, D. S. More about less: A study of language comprehension. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 211-221.

Palermo, D. S. Still more about the comprehension of "less." Developmental Psychology, 1974, 10, 827-829.

Reich, P. A. Notes and discussion: The early acquisition of word meaning. Journal of Child Language, 1976, 3, 117-123.

Weintraub, S. Tables of the Cumulative Binomial Probability Distribution for Small Values of p. New York: Free Press of Glencoe, 1963.

Table 1
Characteristics of the Stimulus Dimensions

Adjective Pair	Feature List	Frequency Estimates			
		+Pol ^a	-Pol ^a	+Pol ^b	-Pol ^b
Natural Concepts:					
big-small	(+Pol)(n-Space)(+Main)	360	542 ^c	223	117 ^d
high-low	(+Pol)(1-Space)(+Vert)(+Top Point)	497	174	210	20
tall-short	(+Pol)(1-Space)(+Vert)	55	212	1	27
wide-narrow	(+Pol)(1-Space)(-Vert)(+Second)	125	63	20	1
deep-shallow	(+Pol)(1-Space)(-Plane)	109	14	13	0
Contrived Concepts:					
H-Surface	(+Pol)(2-Space)(-Vert)				
H-Depth	(+Pol)(1-Space)(-Vert)(-Plane)				
V-Depth	(+Pol)(1-Space)(+Vert)(-Plane)				

^aEstimate from Kučera and Francis (1967).

^bEstimate from Howes (1966).

^cThe word "little" has a frequency count of 430.

^dThe word "little" has a frequency count of 831.

Table 2
Mean Number of Precriterion Errors in the Discrimination Task

		Natural					Contrived		
No. of features		big small 3	tall- short 3	deep- shallow 3	wide- narrow 4	high- low 4	H- surface 3	H- depth 4	V- depth 4
Preschool	+	3.67	3.13	2.83	4.25	2.96	3.54	2.50	2.17
	-	4.42	3.50	2.42	4.58	5.92	3.29	2.92	3.38
	total	4.04	3.31	2.63	4.42	4.44	3.42	2.71	2.77
Adult	+	0.44	0.26	0.13	0.32	0.06	1.06	0.56	0.13
	-	0.75	0.13	0.19	0.38	0.06	0.71	0.63	0.06
	total	0.59	0.19	0.16	0.35	0.00	0.88	0.59	0.09

Table 3
Percent of Correct Responses
To Items on the English Equivalents Test

	big- small	tall- short	deep- shallow	wide- narrow	high- low	H- surface	H- depth	V- depth
Both Poles	80.34	64.27	75.00	55.63	75.59	63.67	53.82	67.71
Positive Pole	75.58	66.65	82.50	59.50	86.30	61.88	59.03	75.69
Negative Pole	85.11	61.89	67.50	51.76	64.87	65.45	48.62	59.73
Difference	-9.53	+4.77	+15.00**	+7.74	+21.43**	-3.58	+10.41*	+15.97**

* $p < 0.05$

** $p < 0.01$

Table 1
Characteristics of the Stimulus Dimensions

Adjective Pair	Feature List	Frequency Estimates			
		+Pol ^a	-Pol ^a	+Pol ^b	-Pol ^b
Natural Concepts:					
big-small	(±Pol)(n-Space)(+Main)	360	542 ^c	223	117 ^d
high-low	(±Pol)(1-Space)(+Vert)(+Top Point)	497	174	210	20
tall-short	(±Pol)(1-Space)(+Vert)	55	212	1	27
wide-narrow	(±Pol)(1-Space)(-Vert)(+Second)	125	63	20	1
deep-shallow	(±Pol)(1-Space)(-Plane)	109	14	13	0
Contrived Concepts:					
H-Surface	(±Pol)(2-Space)(-Vert)				
H-Depth	(±Pol)(1-Space)(-Vert)(-Plane)				
V-Depth	(±Pol)(1-Space)(+Vert)(-Plane)				

^a Estimate from Kučera and Francis (1967).

^b Estimate from Howes (1966).

^c The word "little" has a frequency count of 430.

^d The word "little" has a frequency count of 831.

Table 2
Mean Number of Precriterion Errors in the Discrimination Task

		Natural					Contrived		
		big small 3	tall- short 3	deep- shallow 3	wide- narrow 4	high- low 4	H- surface 3	H- depth 4	V- depth 4
No. of features									
Preschool	+	3.67	3.13	2.83	4.25	2.96	3.54	2.50	2.17
	-	4.42	3.50	2.42	4.58	5.92	3.29	2.92	3.38
	total	4.04	3.31	2.63	4.42	4.44	3.42	2.71	2.77
Adult	+	0.44	0.26	0.13	0.32	0.06	1.06	0.56	0.13
	-	0.75	0.13	0.19	0.38	0.06	0.71	0.63	0.06
	total	0.59	0.19	0.16	0.35	0.00	0.88	0.59	0.09

Table 3
Percent of Correct Responses
To Items on the English Equivalents Test

	big- small	tall- short	deep- shallow	wide- narrow	high- low	H- surface	H- depth	V- depth
Both Poles	80.34	64.27	75.00	55.63	75.59	63.67	53.82	67.71
Positive Pole	75.58	66.65	82.50	59.50	86.30	61.88	59.03	75.69
Negative Pole	85.11	61.89	67.50	51.76	64.87	65.45	48.62	59.73
Difference	-9.53	+4.77	+15.00**	+7.74	+21.43**	-3.58	+10.41*	+15.97**

* $p < 0.05$

** $p < 0.01$